DOI:
10.1016/j.mrgentox.2013.09.007

Document Version
Peer reviewed version

The *GSTM1null* (deletion) and *MGMT84* rs12917 (Phe/Phe) haplotype are associated with bulky DNA adduct levels in human leukocytes

Edith Molina [1], Rebeca Pérez-Morales[2], Julieta Rubio[1], Pavel Petrosyan[1], Leticia Hernández Cadena[3], Volker M. Arlt[4], David H. Phillips[4] and María E. Gonsebatt[1*]

1. Departamento de Medicina Genómica y Toxicología Ambiental, Instituto de Investigaciones Biomédicas, Universidad Nacional Autónoma de México, Apartado Postal 70228, Mexico City 04510, Mexico.
2. Departamento de Biología Molecular, Facultad de Ciencias Químicas, Universidad Juárez del Estado de Durango, Av. Artículo 123 s/n, Col. Filadelfia 35010 Gómez Palacio, Durango, México
3. Centro de Investigación en Salud Poblacional, Instituto Nacional de Salud Pública, Cuernavaca, Morelos, Mexico.
4. Analytical and Environmental Sciences Division, MRC-HPA Centre for Environment and Health, King's College London, Franklin-Wilkins Building, 150 Stamford Street, London SE1 9NH London, United Kingdom.

*Corresponding author. Tel.: (+52) (55) 56229179; Fax: (+52) (55) 56229182; *E-mail address*: margen@.unam.mx

Abstract

Tobacco smoke and air pollutants contain carcinogens, such as polycyclic aromatic hydrocarbons (PAHs) and tobacco specific nitrosamines (TSNA), that are substrates of metabolizing enzymes generating reactive metabolites that can bind to DNA. Variation in the activity of these enzymes may modify the extent to which these metabolites can interact with DNA. We compared the levels of bulky DNA adducts in blood leukocytes from 93 volunteers living in Mexico City with the presence of 13 single nucleotide polymorphisms (SNPs) in genes related to PAH and TSNA metabolism (*AhR* rs2044853, *CYP1A1* rs1048943, *CYP1A1* rs1048943, *CYP1A1* rs1799814, *EPHX1* rs1051740, *EPHX1* rs2234922, *GSTM1* null, *GSTT1* null and *GSTP1* rs947894), DNA repair (*XRCC1* rs25487, *ERCC2* rs13181 and *MGMT* rs12917) and cell cycle (*TP53* rs1042522). ^{32}P-postlabeling analysis was used to quantify bulky DNA adduct formation. Genotyping was performed using PCR-RFLP. The mean levels of bulky DNA adducts were 8.51±3.66 adducts/10^8 nucleotides (nt) in smokers and 8.38±3.59 adducts/10^8 nt in non-smokers, being the difference not statistically significant. Without taking into account the smoking status, *GSTM1null* individuals had a marginally significant lower adduct levels compared with *GSTM1* volunteers (p=0.0433) and individuals heterozygous for *MGMT* Leu/Phe had a higher level of bulky adducts than those who were homozygous wild-type (p=0.0170). A multiple regression analysis model showed a significant association between the *GSTM1* (deletion) and *MGMT* rs12917 (Phe/Phe) haplotype and the formation of DNA adducts in smokers (R^2=0.2401, p=0.0215). The presence of these variants conferred a greater risk for higher adduct levels in this Mexican population.

Keywords

Risk polymorphisms, Gene interaction,^{32}P-postlabeling, *GSTM1*, *MGMT*.

Abbreviations

PAH, polycyclic aromatic hydrocarbon; **SNP,** single nucleotide polymorphism; **mEH,** microsomal epoxide hydrolase; **BPDE,** Benzo[*a*]pyrene-7,8-diol-9,10-epoxide; **BER,** base excision repair; **NER,** nucleotide excision repair; **DSB,** DNA double-strand breaks; **DRR,** direct reversal repair; **ERCC2/XPD,** excision repair cross complementing complementation group 2/ xeroderma pigmentosum D; **XRCC1,** X-ray repair cross complementation-1; **MGMT,** O^6-methylguanine-DNA methyltransferase; **nt,** nucleotides.

1. Introduction

Many polycyclic aromatic hydrocarbons (PAHs), aromatic amines and tobacco specific nitrosamines (TSNA), which are present in tobacco smoke, food and urban air pollution, have been classified as carcinogens by the International Agency for Research on Cancer [1] and the USA Environmental Protection Agency [2]. These compounds have little inherent biological activity, but their toxicity results from the formation of reactive metabolites in the cell and they constitute the major, but not exclusive, source of bulky DNA adducts. Cytochrome 450 (CYP) enzymes play an important role in the generation of these reactive intermediates in a process described as the metabolic activation of PAHs and TSNAs [3, 4]. Experimental evidence has shown that the CYP1 family is modulated by the aryl hydrocarbon receptor (AhR), a ligand-activated transcription factor [5]. Models suggest that CYP1A1 catalyzes the addition of an oxygen atom to lipophilic carcinogens, such as PAHs, thereby increasing their water solubility and producing simple epoxides. Microsomal epoxide hydrolase (mEH) converts epoxides to dihydrodiols, which can subsequently be conjugated with glutathione, catalysed by glutathione S-transferases (GSTs), primarily GST M1, T1 and P1[6]. Thus, soluble metabolites are readily eliminated from organisms. In each of these reactions, highly reactive intermediates are usually formed that can react with DNA and proteins to form covalently bound adducts. Additionally, tobacco and food nitrosamines generate electrophilic metabolites that alkylate DNA [7]. The metabolic activation of nitrosamines to form DNA adducts proceeds via α-hydroxylation (*i.e.*, hydroxylation of the carbon adjacent to the *N*-nitroso group) by CYP enzymes. This α-hydroxylation at the methyl carbon produces α-hydroxymethyl nitrosamines with spontaneous formaldehyde loss and pyridyloxobutyl diazohydroxide production, which

reacts with DNA, thereby forming stable alkyl adducts [8]. Among these adducts, the O^6-methylguanine (O^6-meG) adduct is considered to be one of the most mutagenic DNA lesions, and its carcinogenic effects have been documented. Unrepaired O^6-meG lesions have been associated with increased frequencies of chromosomal aberrations and sister chromatid exchanges [8, 9, 10].

Bulky DNA adducts represent a major and important class of DNA damage originating from cigarette smoke exposure and environmental pollution [11, 12]. DNA adducts are removed by different repair pathways, including base excision repair (BER), nucleotide excision repair (NER) and direct reversal repair (DRR) [11]. MGMT plays an important role in removing primarily the O^6-alkylguanines by DRR. A Leu to Phe change at position 84 in the MGMT protein alters the enzyme substrate affinity [8, 13] and has been associated with increased lung cancer risk, which is greater in female smokers [14].

Similarly, polymorphisms in genes related to PAH and TSNA metabolism, including *AhR*, *CYP1A1*, *EPHX, GSTM1, GSTT1* and *GSTP1* and allelic variants in DNA repair enzymes genes (*i.e.*, *XRCC1, ERCC2* and *MGMT*; Table 1), have been described to contribute to the risk for developing lung cancer, suboptimal repair and PAH and TSNA sensitivity [8, 13, 15].

PAH- and TSNA-DNA adducts in leukocytes are considered biomarkers of exposure to these carcinogens and are therefore cancer-risk indicators [9, 28]. ^{32}P-Postlabeling is a sensitive assay to detect DNA adduct formation and is broadly applicable for a variety of adducts [29]. In the present study we investigated the association between the levels of bulky DNA adducts (measured by ^{32}P-postlabeling) in blood leukocytes from healthy

smokers and non-smokers and the allelic risk variants of genes involved in cigarette smoke carcinogen activation, detoxification and DNA repair in a group of volunteers living in Mexico City.

2. Materials and Methods

2.1.Subjects and DNA isolation

This study included healthy, non-related participants living in Mexico City who donated blood at the "20 de Noviembre" Hospital in Mexico City from October 2001 to November 2004. The study protocol was approved by the Ethics Committee of the Instituto de Investigaciones Biomédicas, Universidad Nacional Autónoma de México, and the "20 de Noviembre" Hospital gave permission to use the buffy coat cells of blood bank samples as a source of DNA. After giving their informed consent, the smoker and non-smoker participants were asked about their overall health status and smoking habits. Smokers with tobacco indices lower than 0.05 or with incomplete data regarding their smoking habits (*i.e.*, the number of cigarettes per day and/or years of smoking) were excluded. For the final analysis, 94 subjects were randomly selected and consisted of 43 smokers and 51 non-smokers. Buffy coat cells were obtained from blood samples and DNA was isolated using a standard phenol-chloroform extraction protocol [30], aliquoted and frozen at –20°C until analysis. The integrity and purity of DNA (0.8 % agarose gel electrophoresis and A260/280 index, respectively) were established for all samples before the ^{32}P-postlabeling assay was initiated in January, 2010.

2.2. ^{32}P-postlabeling

Analysis by ^{32}P-postlabeling was performed according to the previously described standardized procedure [29]. Briefly, DNA samples (4 µg) were digested overnight using a mixture of micrococcal nuclease and spleen phosphodiesterase at 37ºC. Nuclease P_1 digestion was used for adduct enrichment. The labeled DNA adducts were resolved by multi-directional thin-layer chromatography (TLC) on 10 × 20 cm PEI-cellulose plates (Machery-Nagel, Germany). The solvent systems used for TLC were as follows: D1: 1 M sodium phosphate, pH 6.0; D3: 4 M lithium formate, 7 M urea, pH 3.5; and D4: 0.8 M lithium chloride, 0.5 M Tris, 8.5 M urea, pH 8.0. After chromatography, TLC sheets were scanned using a Packard Instant Imager (Dowers Grove, IL, USA), and the DNA adduct levels (RAL; relative adduct labeling) were calculated from the adduct cpm, the specific activity of the [γ-^{32}P]ATP and the amount of DNA (pmol of DNA-P) used. All samples were analyzed in triplicate on different days. An external benzo[*a*]pyrene diol-epoxide-(BPDE)-modified DNA standard was used as a positive control [31]. DNA adduct levels were measured in the diagonal radioactive zone (DRZ) area of the TLC plates. The method provides a summary measure of a complex mixture of adducts present in the postlabeling chromatograms resistant to P1 digestion [32]. The results were expressed as DNA adducts/10^8 nucleotides (nt).

2.3.Genotyping

The polymerase chain reaction (PCR), restriction fragment length polymorphism (RFLP) and TaqMan allelic discrimination assays were performed for genotyping as previously

described by Pérez Morales and coworkers [33]. The results were expressed as DNA adducts/10^8 nucleotides (nt)

2.4.Statistical analysis

Age, cigarettes smoked per day, body mass index (BMI) and bulky DNA adduct levels in leukocytes from smokers and non-smokers were compared by bivariate analysis using the t-test or Wilcoxon signed rank test after testing for normal distribution. The levels of adducts were transformed to Ln for statistical analysis. To explore the relationship between bulky DNA adduct levels and risk polymorphisms, the different allelic variants were grouped by genotypes and scored as follows: 0 (wild type homozygous); 1(heterozygous, one variant allele) and 2 (mutant homozygous). For *GSTM1* and *GSTT1*, the presence of the allele was scored 0, and the null genotype was scored 2 because the presence of *GSTM1* or *GSTT1* in the homozygous or heterozygous condition has not been associated with increased adduct levels [21, 34]. In either case, to evaluate the possible association between adduct levels and each of the polymorphisms, t-test, one way ANOVA or Kruskall-Wallis tests were used accordingly. We controlled for multiple testing by Bonferroni correction using a conservative approach: a priori, we performed the correction for the effective number of independent tests by correlated SNPs, using a principal components analysis [35]. The threshold for array-wide statistical significance was established as $p<0.025$. Separate multivariate models were undertaken for smoker and non-smoker volunteers, these models included selected variables confirmed by the significant association observed in the bivariate analysis including only those that were significant or marginally significant. Finally, the quality of each final model was assessed by testing the statistical assumptions

by linear regression. A distribution of residuals was constructed to evaluate the final multivariate linear regression models. We assessed whether there were any individually influential outliers by residuals plots and Breusch-Pagan/Cook-Weisberg test for heteroscedasticity statistics. A $p<0.05$ was considered statistically significant to include the variables into the final model. All of the statistical analyses were performed using STATA SE.10 (Stata Corporation, College Station, TX) and SigmaPlot 10.0 (Systat Software Inc, San Jose, CA).

All genotypes were in Hardy-Weinberg equilibrium. *GSTM1* and *GSTT1* were not analyzed for Hardy-Weinberg equilibrium because the analytical method used did not allow for the discrimination between heterozygous and homozygous-positive genotypes [33].

3. Results

The bulky adduct levels were analyzed taking into consideration smoking status, age, gender, and BMI (Table 2). Outliers were excluded by principal component analysis. The mean levels of adducts were 8.51±3.66 adducts/10^8 nucleotides (nt) in smokers (n=42) and 8.38±3.59 adducts/10^8 nt in non-smokers (n=51). The data were transformed for statistical analysis; the mean Ln levels of the adducts were 2.05±0.43 adducts/10^8 and 2.04±0.41 adducts/10^8 nt in smokers and non-smokers, respectively, which were not significantly different. On average, smokers had approximately 1.5% more adducts than non-smokers, probably because 88% of the smokers were "low current smokers" (*i.e.*, <10 cigarettes per day). In addition, no significant adduct level differences were observed with respect to gender and BMI.

Thus, the relationship between the bulky DNA adduct levels and the allelic variants were initially analyzed without taking into account the smoking status (Table 3). Individuals with *GSTM1null* genotype had a marginally significant lower adduct levels compared with those with *GSTM1* genotype (p=**0.0433**). Moreover, those heterozygous for *MGMT* Leu/Phe had a higher level of bulky adducts (2.240 ±0.448 adducts/10^8 nt) than those who were homozygous wild-type (1.963±0.364 adducts/10^8 nt) (**p=0.0170**).

If smoking status was taken into consideration, smokers with *XRCC1* Arg/Gln, *ERCC2* Lys/Gln or *MGMT84* Leu/Phe genotypes showed higher levels of bulky adducts (2.24±0.39 adducts/10^8 nt, p=0.023; 2.27±0.51, p=0.16 and 2.240±0.448 adducts/10^8 nt, p=0.256, respectively) than those who were homozygous wild-type. Between non-smokers, only those with *AhR* Arg/Lys had lower levels of bulky adduct (p=0.045) compared with homozygous wild-type (data not shown).

We performed a regression analysis model to associate the adduct levels versus the allelic variants through interactions with different genetic polymorphisms and their possible combinations in 38 smokers and 43 non-smokers (Table 4). Significant predictor haplotypes were found among non-smokers: *AhR* rs2066853 (Arg/Lys) was positively and significantly associated with adduct formation (β=0.327; p=0.006), and *GSTM1null* (deletion) and *MGMT84* rs12917 (Phe/Phe) were negatively associated with bulky DNA adduct formation (β=-0.641; p=0.035). These haplotypes explain 10 and almost 7% of the variability, respectively. Among smokers, the interaction between *GSTM1null* (deletion) and *MGMT84* rs12917 (Phe/Phe) was positively and significantly associated with adduct formation (β=1.055; p=0.015), explaining 14% of the variability. To validate the

multivariate linear regression model, we examined the residuals. Across all of the predicted bulky adducts, the level of the distribution of the residuals was normal and homoscedastic (Chi^2 (1)=0.45; Probability > Chi^2=0.5031). Thus, the *AhR* rs2066853 (Arg/Lys), *GSTM1null* (deletion) and *MGMT84* rs12917 (Phe/Phe) haplotype represents our predictive model for the higher risk of bulky adduct formation in this Mexican population.

4. Discussion

Molecular epidemiology studies have shown significant interindividual variations in the levels of DNA adduct formation [32]. These variations have been attributed to the differences in the metabolism and elimination of xenobiotics and DNA repair processes, which could be explained by the presence of genetic polymorphisms in a proportion of the genes encoding the enzymes that participate in these processes. Moreover, several epidemiological studies have shown that single nucleotide polymorphisms (SNPs) are associated with differential biological activities or increased cancer risk (Table 1). We investigated the relationship between the bulky DNA adduct levels detected by ^{32}P-postlabeling assay in leukocytes from healthy smokers and non-smokers and 13 SNPs involved in PAH and TSNA metabolism (*AhR (rs2066853)*, *CYP1A1 (rs4646903)*, *CYP1A1 (rs1048943)*, *CYP1A1* (rs1799814), *EPHX1* (rs1051740), *EPHX1* (rs2234922), *GSTM1*, *GSTT1* and *GSTP1* (rs947894)), DNA repair (*XRCC1* (rs25487), *ERCC2* (rs1318) and *MGMT84* (rs12917)) and the cell cycle (*TP53* (rs1042522)).

No association was found between age and adduct levels (Table 2). Positive correlations between the smoking-related adduct levels and age have been reported only in persons above the age of 60, presumably reflecting environmental, lifestyle, DNA repair and genetic factors [36, 37]. No differences were found between bulky adduct levels in smokers and non-smokers, probably because 88% (38 out of 42) of the smokers were “low current” smokers according to the number of cigarettes smoked per day (*i.e.*, <10 cigarettes/day), and no correlations were found between PAH-related DNA adduct formation and the low current smoking status in similar studies on Caucasians [38]. Alternatively, diet is also a potential source for DNA adduct formation in blood [38, 39, 40]. The levels of adducts can

be altered as a function of the quantity of PAHs ingested during consumption of charcoal-broiled beef [41]; however, we have no information on the diet composition of these volunteers to allow similar analysis.

Individuals with *GSTM1null* genotype had a marginally significantly lower adduct levels compared with those who were *GSTM1* positive (p=**0.0433**; Table 3). Higher PAH-DNA adduct levels have been associated to *GSTM1null* genotype in Caucasians; in contrast, no differences were found in African-American and Latino populations with respect to the *GSTM1* status [42]. There are a great number of studies on *GSTM1null* individuals which report different relations of this null genotype and DNA adduct levels quantified by different techniques [42, 43, 44]. A limitation of our study was the relatively small number of *GSTM1null* subjects. To clarify the participation of *GSTM1 null* genotype in the Mexican population a larger sample needs to be analyzed in future studies.

In human lymphocytes, the *MGMT* rs12917 (Leu84Phe) polymorphism results in suboptimal repair of genetic damage and increased NNK (4-[methylnitrosamino]-1-[3-pyridyl]-1-butanone) sensitivity, as detected by the formation of *in vitro* chromosomal aberrations [8]. Moreover, the *MGMT* rs12917 (Leu84Phe) polymorphism and suboptimal repair were also associated with increased mutation frequency in the lymphocytes of smokers compared with smokers carrying the wild-type allele [10, 44]. Accordingly, we observed significantly higher adduct levels in the individuals heterozygous for *MGMT* Leu/Phe than in the homozygous wild-type individuals (**p=0.0170**; Table 3). These results are interesting because the typical adducts which are repaired by MGMT enzyme involve alkyl and POB adducts derived from NNK metabolisms [45] that are generally detected by

HPLC-immunoassay or by enrichment with immunoaffinity columns prior to ^{32}P-postlabeling [46, 47], procedures that were not used in this study.

The conditions used in this ^{32}P-postlabeling assay allows for quantification of a wide range of bulky DNA adducts [48, 49, 50] which are revealed in the diagonal radioactivity zone (DRZ) on 2-dimensional thin-layer chromatograms generated by resolution of the labeled digests. Complex mixture like tobacco smoke and air pollution contain over 4000 compounds and can potentially generate numerous unidentified DNA adducts in the tissues [51]. Using urea based solvents, detection in the DRZ of intrastrand DNA-DNA cross links formed by treatment of DNA with hydroxyl radicals has been reported [52, 53] suggesting that some of the DNA adducts detectable in the DRZ in our assays may also represent DNA cross-links at the O^6 position of guanine generated by butadiene, catechols and aldehydes presents in tobacco smoke [54, 55]. The human MGMT enzyme could participate in the repair of such products [56].

Another hypothesis on the relationship of MGMT and bulky DNA adduct formation is that MGMT might bind directly to AhR and prevent AhR association with its co-activators, which down-regulates the AhR-mediated transcription of the genes implicated in the activation of the xenobiotics. It was suggested that MGMT could serve as a sensor of DNA exposure to alkylating agents (which are, in our case, concomitant to the typical bulky xenobiotic AhR ligands occurring in air and smoke pollution) thus controlling nuclear receptors implied in cell proliferation according to the extent of DNA damage [57, 58, 59]. As yet, no molecular interaction between AhR and MGMT has been documented; however, the fact that MGMT contains the consensus LxxLL motif of transcriptional co-activators

[60, 61] fully recognized by AhR [62] makes this hypothesis plausible. Indeed, MGMT was shown to bind to estrogen receptor (ER) through its LxxLL motif-containing peptide domain exposed upon suicide DNA repair trans-alkylation of MGMT [63]. This binding modulates ER association with ER co-activators to negatively regulate ER-mediated transcription [64]. Based on the data obtained in this study on the relationship of MGMT polymorphism and bulky DNA adduct formation, the alkylated Leu84Phe allelic variant MGMT is suggested to have decreased affinity for AhR binding and to exert diminished AhR sequestering activity compared to the alkylated form of wild-type MGMT. As a consequence, more AhR would be available for PAH binding in MGMT Leu84Phe cells resulting in effective transcription activation of, e.g., phase I enzymes CYP1A1, CYP1A2 that activate PAH-DNA binding. This idea is supported by the notion that Leu84Phe substitution is adjacent to His85 residue involved Zn (II) binding in MGMT [59] and is separated by 15 aa from the LxxLL motif.

The individual genetic risk from exposure to tobacco carcinogens is governed by a combination of genetic polymorphisms in different xenobiotic-metabolizing genes. Our model suggests that individuals with the *AhR* (rs2066853), *GSTM1null* (deletion) and the *MGMT 84* (rs12917) haplotypes exhibit an increased risk for the formation of bulky DNA adducts in leukocytes. This model explains 24% of the DNA adducts in leukocytes, while the 76% of the variability remains unexplained (Table 4). The strength of this model is that we selected only risk haplotypes associated with the higher adduct levels. It is also important to consider that we used a simple regression model to account for interacting gene effects; however, these observations require confirmation from in a larger study. This model assumes that the three genes have a meaningful biological relationship that is linear

and continuous, and the risk may be enhanced by exposure to environmental PAHs~~,~~ or tobacco-smoke carcinogens.

Funding and Acknowledgments

This work was supported by grant from CONACYT 46341-M (M.E.G.) and grant from the Universidad Nacional Autónoma de México PAPIIT IN219807 (J.R). Edith Molina received a scholarship from CONACYT (No. 199980) and is a graduate student at Doctorado en Ciencias Biomédicas, UNAM. Volker M. Arlt and David H. Phillips are supported by Cancer Research UK and are also members of ECNIS2 (Environmental Cancer, Nutrition and Individual Susceptibility) European Union Network of Excellence. We thank M.C. Clementina Castro-Hernández for helping us with blood bank samples.

Conflicts of interest

None declared

References

1. International Agency for Research on Cancer (IARC) Monographs on the Evaluation of Carcinogenic Risks to Humans. IARC, Lyon, France. (2004) Volume 83.
2. Environmental Protection Agency (EPA) (2006). Integrated Risk Information System (IRIS). (Accessed February 26, 2010). Available from *URL: www.epa.gov/ncea/iris*
3. J. H. Kim, M. E. Sherman, F.C. Curriero, F.P. Guengerich, P. T. Strickland, , T. R. Sutter, Expression of cytochromes P450 1A1 and 1B1 in human lung from smokers, non-smokers, and ex-smoker, Toxicology and Applied Pharmacology, 99 (2004) 210-219.
4. R. Nilson, The molecular basis for induction of human cancers by tobacco specific nitrosamines, Regul Toxicol Pharmacol. 60 (2011) 268-280.

5. S. R. Kondraganti, P. Fernandez-Salguero, F.J. Gonzalez, K.S. Ramos, W. Jiang, B. Moorthy, Polycyclic aromatic hydrocarbon-inducible DNA adducts: evidence by 32P-postlabeling and use of knockout mice for Ah receptor-independent mechanisms of metabolic activation in vivo, Int J Cancer. 103 (2003) 5-11.
6. O. Pelkonen, D. W. Nebert, Metabolism of polycyclic aromatic hydrocarbons: etiologic role in carcinogenesis. Pharmacol Rev. 34 (1982) 189-222.
7. B. Kaina, M. Christmann, S. Naumann, W.P. Roos, MGMT: key node in the battle against genotoxicity, carcinogenicity and apoptosis induced by alkylating agents, DNA Repair (Amst).6 (2007) 1079-1099.
8. C. E. Hill, J.K. Wickliffe, K.J. Wolfe, C.J. Kinslow, M.S. Lopez, S.Z. Abdel-Rahman, The L84F and the I143V polymorphisms in the O6-methylguanine-DNA-methyltransferase (MGMT) gene increase human sensitivity to the genotoxic effects of the tobacco-specific nitrosamine carcinogen NNK, Pharmacogenet Genomics. 15 (2005) 571-578.
9. G.P. Margison, M.F. Santibanez Koref, A.C. Povey, Mechanisms of carcinogenicity/chemotherapy by O6-methylguanine, Mutagenesis. 17 (2002) 483-487.
10. C. E. Hill, J.K. Wickliffe, A.T. Guerin, C.J. Kinslow, K.J. Wolfe, M.M. Ammenheuser, S.Z. Abdel-Rahman, The L84F polymorphism in the O6-Methylguanine-DNA-Methyltransferase (MGMT) gene is associated with increased hypoxanthine phosphoribosyltransferase (HPRT) mutant frequency in lymphocytes of tobacco smokers, Pharmacogenet. Genomics. 17 (2007) 743-753.
11. B. Hang, Formation and repair of tobacco carcinogen-derived bulky DNA adducts, J Nucleic Acids. Dec 20 (2010) 709521.
12. P. Rossner, V. Svecova, J. Schmuczerova, A. Milcova, N. Tabashidze, J. Topinka, A. Pastorkova, R.J. Sram, Analysis of biomarkers in a Czech population exposed to heavy air pollution. Part I: bulky DNA adducts, Mutagenesis. Oct 9. (2012) [Epub ahead of print].
13. A. C. Povey, G.P. Margison, M.F. Santibáñez-Koref, Lung cancer risk and variation in MGMT activity and sequence, DNA Repair (Amst). 6 (2007) 1134-1144.
14. L. Wang, H. Liu, Z. Zhang, M.R. Spitz, Q. Wei, Association of genetic variants of O6-methylguanine-DNA methyltransferase with risk of lung cancer in non-Hispanic Whites, Cancer Epidemiol Biomarkers Prev. (2006), 15, 2364-2369.
15. C. Graziano, C.E. Comin, C. Crisci, L. Novelli, L. Politi, L. Messerini, M. Andreani, B. Porfirio, Functional polymorphisms of the microsomal epoxide hydrolase gene: a reappraisal on a early-onset lung cancer patients series, Lung Cancer. 63(2009) 187-193.
16. J. M. Wong, A. B. Okey, P.A. Harper, Human aryl hydrocarbon receptor polymorphisms that result in loss of CYP1A1 induction. Biochem Biophys Res Commun. 288 (2001): 990-996.

17. D. Petersen, C. McKinney, K. Ikeya, H. Smith, A. Bale, O. Wesley, D. Nebert,. Human CYP1A1 Gene: Cosegregation of Enzyme Inducibility Phenotype and an RFLP. Am. J. Hum. Genet. 48(1991):720-725.
18. D. Schwarz, P. Kisselev, A. Chernogolov, W. Schunck, I. Roots,. Human CYP1A1 variants lead to differential eicosapentaenoic acid metabolite patterns. Biochemical and Biophysical Research Communications. 336(2005):779-783
19. M.P. Gallegos, L.E. Figuera, R. Troyo, G. Morgan, A.M. Puebla, M.R. Flores, G.M Zuñiga. CYP1A1 *2B and *4 polymorphisms are associated with lung cancer susceptibility in Mexican patients. Inter J Biol Markers. 23 (2008): 24-30.
20. M. Lodovici, C. Luceri, F. Guglielmi, C. Bacci, V. Akpan, M.L. Fonnesu, V. Boddi, P. Dolara, Benzo(a)pyrene diolepoxide (BPDE)-DNA adduct levels in leukocytes of smokers in relation to polymorphism of CYP1A1, GSTM1, GSTP1, GSTT1, and mEH. Cancer Epidemiol Biomarkers Prev.13 (2004) 1342-8.
21. M. Rojas, I. Cascorbi, K. Alexandrov, E. Kriek, G. Auburtin, L. Mayer, A. Kopp-Schneider, I. Roots, H. Bartsch, Modulation of benzo[a]pyrene diolepoxide-DNA adduct levels in human white blood cells by CYP1A1, GSTM1 and GSTT1 polymorphism, Carcinogenesis. 21(2000) 35-41.
22. S. Pavanello, A. Pulliero, E. Siwinska, D. Mielzynska, E.Clonfero, Reduced nucleotide excision repair and GSTM1-null genotypes influence anti-B[a]PDE-DNA adduct levels in mononuclear white blood cells of highly PAH-exposed coke oven workers. Carcinogenesis, 26 (2005) 169-175.
23. D. Ryberg, V. Skaug, A. Hewer, D.H. Phillips, L.W. Harries, C.R. Wolf, D. Ogreid, A. Ulvik, P, Vu, A. Haugen, Genotypes of glutathione transferase M1 and P1 and their significance for lung DNA adduct levels and cancer risk. Carcinogenesis. 18(1997):1285-9.
24. S. Zienolddiny, D. Campa, H. Lind, D. Ryberg, V. Skaug, L. Stangeland, D.H. Phillips, F. Canzian, A. Haugen, Polymorphisms of DNA repair genes and risk of non-small cell lung cancer. Carcinogenesis, 27(2006), 560-567.
25. R. M. Lunn, K.J. Helzlsouer, R. Parshad, D.M. Umbach, E.L. Harris, K.K. Sanford, D.A. Bell, XPD polymorphisms: effects on DNA repair proficiency. Carcinogenesis, 21, (2000)551-555.
26. H. Zhao, L.E. Wang, D. Li, R.M. Chamberlain, E.M. Sturgis, Q. Wei, Genotypes and haplotypes of ERCC1 and ERCC2/XPD genes predict levels of benzo[a]pyrene diol epoxide-induced DNA adducts in cultured primary lymphocytes from healthy individuals: a genotype-phenotype correlation analysis. Carcinogenesis, 29, (2008) 1560-1566.
27. C. E. Irarrazabal, C. Rojas, R. Aracena, C. Márquez, L. Gil, Chilean pilot study on the risk of lung cancer associated with codon 72 polymorphism in the gene of protein p53. Toxicol Lett, 144 (2003) 69-76.

28. M. Poirier, Chemical-induced DNA damage and human cancer risk. Nature. Aug; 4(2004) 630-7. Review.
29. D. H. Phillips, V. M. Arlt, The 32P-postlabeling assay for DNA adducts. Nat Protoc. 2(2007): 2772-2781.
30. J. Sambrook, E. F. Fritsch, T. Maniatis, Molecular Cloning: A Laboratory Manual, 2nd ed. Cold Spring Harbor Laboratory Press, New York, (1989), E3–E4.
31. D.H. Phillips, M. Castegnaro, Standardization and validation of DNA adduct postlabelling methods: report of interlaboratory trials and production of recommended protocols, Mutagenesis, 1999. 14(3):301-15.
32. D. Tang, D.H. Phillips, M. Stampfer, L. Mooney, Y. Hsu, Y. Cho, W-Y. Tsai, J. Ma, K. Cole, M. Ni She´, F Perera, Association between Carcinogen-DNA Adducts in White Blood Cells and Lung Cancer Risk in the Physicians Health Study, Cancer Research 61(2001): 6708–6712.
33. R. Pérez-Morales, I. Méndez-Ramírez, C. Castro-Hernández, O.C. Martínez-Ramírez, M.E. Gonsebatt, J. Rubio, Polymorphisms associated with the risk of lung cancer in a healthy Mexican Mestizo population: Application of the additive model for cancer, Genet Mol Biol. 34(2011): 546-552.
34. A. Rundle, D. Tang, J. Zhou, S. Cho, F. Perera, The association between glutathione S-transferase M1 genotype and polycyclic aromatic hydrocarbon-DNA adducts in breast tissue, Cancer Epidemiol Biomarkers Prev. 9(2000)1079-85.
35. X. Gao, J. Starmer, E.R. Martin , A multiple testing correction method for genetic association studies using correlated single nucleotide polymorphisms. Genetic Epidemiol. 32(2008): 361-369.
36. K. Randerath, K.L. Putman, H.H. Osterburg, S.A. Johnson, D.G. Morgan, C.E. Finch, Age-dependent increases of DNA adducts (I-compounds) in human and rat brain DNA. Mutat Res. 295(1993): 11-18.
37. F. Ricceri, R.W. Godschalk, M. Peluso, D.H. Phillips, A. Agudo, P. Georgiadis, S. Loft, A. Tjonneland, O. Raaschou-Nielsen, D. Palli, F. Perera, R. Vermeulen, E. Taioli, R.J. Sram, A. Munnia, F. Rosa, A. Allione, G. Matullo, P. Vineis, Bulky DNA adducts in white blood cells: a pooled analysis of 3,600 subjects, Cancer Epidemiol Biomarkers Prev. Dec 19(2010) 3174-81.
38. M.M. Pratt, K. John, A.B. MacLean, S. Afework, D.H. Phillips, M.C. Poirier, Polycyclic aromatic hydrocarbon (PAH) exposure and DNA adduct semi-quantitation in archived human tissues, Int J Environ Res Public Health 8(2011): 2675-2691.
39. D. Palli, P. Vineis, A. Russo, F. Berrino, V. Krogh, G. Masala, A. Munnia, S. Panico, E. Taioli, R. Tumino, S. Garte, M. Peluso, Diet, metabolic polymorphisms and DNA adducts: The Epic- Italy cross-sectional study, Int. J. Cancer. 87 (2000) 444-451.
40. M. Peluso, L. Airoldi, A. Munnia, A. Colombi, F. Veglia, H. Autrup, A. Dunning, S. Garte, E. Gormally, C. Malaveille, G. Matullo, K. Overvad, O. Raaschou-

Nielsen, F. Clavel-Chapelon, J. Linseisen, H. Boeing, A. Trichopoulou, D. Palli, V. Krogh, R. Tumino, S. Panico, B.H. Bueno-De-Mesquita, P.H. Peeters, M. Kumle, A. Agudo, C. Martinez, M. Dorronsoro, A. Barricarte, M.J. Tormo, J.R. Quiros, G. Berglund, B. Jarvholm, N.E. Day, T.J. Key, R. Saracci, R. Kaaks, E. Riboli, S. Bingham, P, Vineis, Bulky DNA adducts, 4-aminobiphenyl-haemoglobin adducts and diet in the European Prospective Investigation into Cancer and Nutrition (EPIC) prospective study, Br. J. Nutr. 100(2008):489-95.

41. N. Rothman, M.C. Poirier, M.E. Baser, J.A. Hansen, C. Gentile, E.D. Bowman, P.T. Strickland, Formation of polycyclic aromatic hydrocarbon-DNA adducts in peripheral white blood cells during consumption of charcoal- broiled beef. Carcinogenesis.11 (1990)1241-1243.
42. K.F. Weiserbs, J.S. Jacobson, M.D. Begg, L.W. Wang, Q. Wang, M. Agrawal, E.P. Norkus, T.L. Young, R.M. Santella, A cross-sectional study of polycyclic aromatic hydrocarbon-DNA adducts and polymorphism of glutathione S-transferases among heavy smokers by race/ethnicity, Biomarkers (2003), 8:142-55.
43. R. Piipari, T. Nurminen, K. Savela, A. Hirvonen, T. Mantyla, S. Anttila, Glutathione S-transferases and aromatic DNA adducts in smokers bronchoalveolar macrophages, Lung Cancer;39 (2003)265-72.
44. M. Christmann, B. Kaina, O(6)-methylguanine-DNA methyltransferase (MGMT): impact on cancer risk in response to tobacco smoke, Mutat. Res.736(2012) 64-74. Epub 2011 Jun 14.
45. S.S Hecht, Biochemistry, biology and carcinogenicity of Tobacco-Specific *N*-Nitrosamines, Chemical Research in Toxicology. 11(1998) 560-590.
46. K. Randerath, E. Randerath, H.P. Agrawal, R.C. Gupta, M.E. Schurdak, R. Vijayaraj, Postlabeling Methods for Carcinogen-DNA adduct analysis, Environmental Health Perspectives. 62(1985) 57-65.
47. P. Vodicka, R. Stetina, R. Kumar, K. Plna, K. Hemminki, 7-Alkylguanine adducts of styrene oxide determined by 32P-postlabelling in DNA and human embryonal lung fibroblast (HEL), Carcinogenesis (1996), 17, 801-808.
48. R.C. Gupta, M. Vijayaraj Reddy, K. Randerath, 32P-postlabeling analysis of non-radioactive aromatic carcinogen-DNA adducts, Carcinogenesis. 3(1982) 1081-1092.
49. D.H. Phillips, A. Hewer, P.L. Grover, Aromatic DNA adducts in human bone marrow and peripheral blood leukocytes, Carcinogenesis. 7 (1986) 2071–2075.
50. V.M. Arlt, C.A. Bieler, W. Mier, M. Wiessler, H.H. Schmeiser, DNA adduct formation by the ubiquitous environmental contaminant 3-nitrobenzanthrone in rats determined by 32P-postlabeling, Int. J. Cancer 93 (2001) 450–454.
51. D. Hoffmann, I. Hoffmann, K. El-Bayoumy, The less harmful cigarette: a controversial issue. A tribute to Ernst L. Wynder, Chem. Res. Toxicol. 14 (2001) 767-790.

52. K. Randerath, E. Randerath, C. V. Smith, J. Chang, Structural origins of bulky oxidative DNA adducts (type II I-compounds) as deduced by oxidation of oligonucleotides of known sequence, Chem. Res. Toxicol. 9 (1996) 247-254.
53. D. R. Lloyd, D. H. Phillips, P. L. Carmichael, Generation of putative intrastrand cross-links and strand breaks in DNA by transition metal ion-mediated oxygen radical attack. Chem. Res. Toxicol. 10 (1997) 393-400.
54. . I. Seeman, M. Dixon, H-J. Haussmann, Acetaldehyde in Mainstream Tobacco Smoke: Formation and Occurrence in Smoke and Bioavailability in the Smoker, Chemical Research in Toxicology, 15(2002) 1332-1346.
55. J.M. Arif, C. Dresler, M.L. Clapper, C.G. Gairola, C. Srinivasan, R.A. Lubet, R.C. Gupta, Lung DNA adducts detected in human smokers are unrelated to typical polyaromatic carcinogens, Chem Res Toxicol.19 (2006) 295-299.
56. J.L. Tubbs, A.E. Pegg, J.A. Tainer, DNA binding, nucleotide flipping, and the helix-turn-helix motif in base repair by O6-alkylguanine-DNA alkyltransferase and its implications for cancer chemotherapy, DNA Repair (Amst), (2007) 1100-15.
57. T. C. Ayi, H. K. Oh., T. K. Lee, B. F. Li, A method for the simultaneous identification of human active and active-site alkylated DNA repair enzyme, O6 - methylguanine-DNA methyltransferase, and its application for monitoring human exposure to alkylating carcinogens, Cancer Res.54 (1994) 3726–3731.
58. R.B. Ali, A.K. Teo, H.K. Oh, L.S. Chuang, T.C. Ayi, B.F. Li, Implication of localization of human DNA repair enzyme O6-methylguanine-DNA methyltransferase at active transcription sites in transcription-repair coupling of the mutagenic O6-methylguanine lesion, Mol. Cell. Biol.8 (1998) 1660–1669.
59. Y. Mishina, E.M. Duguid, C. He, Direct reversal of DNA alkylation damage. Chem Rev.106 (2006) 215-32.
60. D.M. Heery, E. Kalkhoven, S. Hoare, M.G. Parker, A signature motif in transcriptional co-activators mediates binding to nuclear receptors, Nature. 12 (1997) 733-736.
61. S. Westin, R. Kurokawa, R.T. Nolte, G.B. Wisely, E.M. Mclnerney, D.W. Rose, M.V. Milburn, M.G. Rosenfeld, C.K. Glass, Interactions controlling the assembly of nuclear receptor heterodimers and co-activators, Nature 395, (1998)199-202.
62. C. Flaveny, R.K. Reen, A. Kusnadi, G.H. Perdew, The mouse and human Ah receptor differ in recognition of LXXLL motifs. Arch Biochem Biophys, 471 (2008) 215-223.
63. H.K. Oh, A.K. Teo, R.B. Ali, A. Lim, T.C. Ayi, D.B. Yarosh, B.F. Li, Conformational change in human DNA repair enzyme O6-methylguanine-DNA methyltransferase upon alkylation of its active site by SN1 (indirect-acting) and SN2 (direct-acting) alkylating agents: breaking a "salt-link", Biochemistry. 35(1996):12259-12266.
64. A.K. Teo, H.K. Oh, R.B. Ali, B.F. Li, The modified human DNA repair enzyme O(6)-methylguanine-DNA methyltransferase is a negative regulator of estrogen

receptor-mediated transcription upon alkylation DNA damage, Mol Cell Biol. (2001) Oct; 21, 7105-7114.

Table 1. Alellic variants in genes associated with increased health risk due to exposure to PAHs and TSNAs and their biological effect.

Gen	SNP	Biological Effect	References
AhR (rs2066853)	Arg554Lis	Expression of CYP1 family	[16]
CYP1A1 (rs4646903)	T3801C	Higher CYP1A1 expression	[17]
CYP1A1 (rs1048943)	Ile462Val	Higher CYP1A1 activity	[18]
CYP1A1 (rs1799814)	Thr461Asn	Lower CYP1A1 activity	[19]
EPHX1 (rs1051740)	Tyr113His	50% activity decrease	[20]
EPHX1 (rs2234922)	His139Arg	25% activity increase	
GSTM1 (deletion)	null	Higher DNA adduct levels	[21]
GSTT1 (deletion)	null		[22]
GSTP1 (rs947894)	Ile105Val		[23]
XRCC1 (rs25487)	Arg399Gln	Decrease in lung cancer risk	[24]
ERCC2 (rs1318)	Lys751Gln	Higher DNA adduct levels	[25], [26]
MGMT84 (rs12917)	Leu84Phe	Suboptimal repair and human sensitivity to TSNA	[8], [13]
TP53 (rs1042522)	Arg72Pro	Increases the risk of cancer, COPD and bronchial diseases	[27]

Table 2. Age, cigarettes smoked per day, BMI and bulky DNA adduct levels in leukocytes from smokers and non-smokers.

	N	Non smokers mean ±S.D.	N	Smokers mean ±S.D.	*P value*
Age [years]	51	36.92±10.92	42	36.37±10.85	
Female	26	38.96±10.97	20	34.75±9.296	NS
Male	25	34.80±9.971	22	37.78±12.08	
Number of cigarettes smoked per day	51	0	42	5.99±5.95	< 0.0001[a]
Female	26	0	20	5.92±4.92	< 0.0001[a]
Male	25	0	22	6.25±6.93	< 0.0001[a]
Ln (adducts/ 10^8nt)	51	2.04±0.41	42	2.05±0.43	
Female	26	2.01±0.42	20	2.12±0.39	NS
Male	25	2.07±0.39	22	1.99±0.46	
BMI	50[b]	26.98±4.70	42	26.2±3.28	NS
Female	26	27.95±4.79	20	26.65±2.97	

Male	24[b]	25.93±4.46	22	25.80±3.56	

Body mass index (BMI)

[a] Wilcoxon signed rank test

[b] One sample without BMI

NS No significant differences in each group ($p>0.05$) using Student's *t*-test.

Table 3. Mean Ln of adducts /10^8 nt and Median of adducts/10^8 nt associated with the different allelic variants.

Genotypes	N[c]	Ln of adducts /10^8 nt	Adducts /10^8 nt	*P* value*
		Mean (S.D)	Median (IQR)	
AhR (rs2066853)				
Arg/Arg	52	2.056 (0.383)	7.69 (3.98)	0.1808[a]
Arg/Lys	29	1.931 (0.462)	5.98 (3.86)	
Lys/Lys	0	-	-	
CYP1A1 (rs1048943)				
Ile/Ile	17	1.917 (0.413)	6.73 (4.83)	0.3197[b]
Ile/Val	55	2.061 (0.387)	7.55 (4.69)	
Val/Val	21	2.116 (0.479)	8.48 (7.12)	
CYP1A1 (rs4646903)				
T/T	24	2.038 (0.475)	7.26 (6.56)	0.9757[b]
T/C	49	2.044(0.339)	7.83 (3.60)	

C/C	20	2.065 (0.518)	7.58 (7.76)	
CYP1A1 (rs1799814)				
Thr/Thr	78	2.063 (0.419)	7.76 (4.79)	0.6485[c]
Thr/Asn	13	1.947 (0.405)	7.14 (3.82)	
Asn/Asn	2	2.076 (0.435)	8.36 (4.99)	
EPHX1 (rs1051740)				
Tyr/Tyr	23	2.146 (0.447)	9.19 (6.14)	0.1133[b]
Tyr/His	41	2.082 (0.411)	7.83 (4.75)	
His/His	29	1.919 (0.374)	6.39 (4.22)	
EPHX1 (rs2234922)				
His/His	70	2.021 (0.419)	7.38 (4.79)	0.2973[a]
His/Arg	23	2.126 (0.397)	8.48 (4.86)	
Arg/Arg	0	-	-	
GSTM1 (deletion)				
Positive	63	2.106 (0.411)	8.50 (4.75)	**0.0433[a]**

Null	30	1.922 (0.401)	6.54 (4.66)	
GSTT1 (deletion)				
Positive	74	2.025 (0.428)	7.43 (4.89)	0.3143[a]
Null	19	2.133 (0.354)	8.75 (4.5)	
GSTP1 (rs947894)				
Val/Val	19	2.083 (0.445)	7.51 (5.82)	0.5935[b]
Ile/Val	51	2.068 (0.411)	8.10 (4.70)	
Ile/Ile	23	1.971 (0.406)	6.22 (5.24)	
XRCC1 (rs25487)				
Arg/Arg	53	1.992 (0.430)	6.59 (4.26)	0.2091[c]
Arg/Gln	36	2.123 (0.394)	8.59 (5.38)	
Gln/Gln	2	1.816 (0.258)	6.25 (2.26)	
ERCC2 (rs1318)				
Lys/Lys	59	2.060 (0.393)	7.83 (4.75)	0.9259[b]
Lys/Gln	27	2.028 (0.457)	7.55 (5.21)	

Gln/Gln	7	2.015 (0.483)	5.98 (5.73)	
MGMT (rs12917)				
Leu/Leu	54	1.963 (0.364)	7.26 (3.82)	**0.0170**[b]
Leu/Phe	26	2.240 (0.448)	10.16 (6.54)	
Phe/Phe	13	2.012 (0.447)	7.37 (4.78)	
TP53 (rs1042522)				
Arg/Arg	47	2.074 (0.422)	7.69 (4.77)	0.4756[b]
Arg/Pro	31	1.984 (0.428)	7.38 (4.84)	
Pro/Pro	15	2.095 (0.371)	8.75 (5.01)	

[a] *t*-test with equal variances for two groups

[b] One way ANOVA test with equal variances

[c] Kruscall-Wallis test (non parametric)

[d] Numbers do not add up to expected totals because of missing genotyping data.

[*] p value < 0.025 was considered significant after Bonferroni correction using correlated SNPs.

Table 4. Multiple linear regression analysis of predictors for bulky DNA adducts.

Covariates	Non-smokers (N=43)[e]				Smokers (N=38)[e]			
	Multivariate			r^2	Multivariate			r^2
	β [a] Coefficient	*P* [b] value	Prob > F [c]	variability explained [d]	β [a] Coefficient	*P* [b] value	Prob > F [c]	variability explained [d]
***AhR* (Arg/Lys)**	**0.327**	**0.006**	**0.0365**	10.67%	-0.0123	0.931	**0.0215**	0.64%
***XRCC1* (Arg/Gln)**	0.066	0.601		0.3%	-0.204	0.140		1.34%
***ERCC2/XPD* (Lys/Gln)**	-0.169	0.153		2.52%	0.037	0.822		0.012%
Interaction of *GSTM1 null* with *MGMT* (Phe/Phe)	**-0.641**	**0.035**		6.94%	**1.055**	**0.015**		14.55%
r^2	18.17%				24.01%			

[a] β Regression coefficient: the unit change of dependent variable caused by 1 unit change of independent variable. For genotype covariates, β is the difference of the mean of each genotype coded as 0 or 1.